Irish Interdisciplinary Journal of Science & Research (IIJSR)
Volume 10, Issue 2, Pages 12-22, April-June 2026

A Brief Survey on RL-Based Adaptive Defender against Advanced Multi-Path Threats

Regavarsha S.[1*], Ananth Kumar T.[2] & Kanimozhi P.[3]

[1,2,3]*Department of Computer Science of Engineering, IFET College of Engineering, Villupuram, Tamil Nadu, India.*
Corresponding Author (Regavarsha S.) Email: regavarsha49@gmail.com[*]

DOI: https://doi.org/10.46759/iijsr.2026.10202

Copyright © 2026 Regavarsha S. et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Article Received: 21 February 2026 *Article Accepted: 25 April 2026* *Article Published: 27 April 2026*

ABSTRACT

One issue that is prevalent in cybersecurity is Advanced Persistent Threat. They occur in a series of operations, and remain undetected over a considerable time, and are simply inconspicuous to normal security devices. The current paper considers one method of counter fighting by means of reinforcement learning. We have reviewed some of the latest studies and also examined closely one of the suggested systems that use simulation as a method of training an AI defender. It is supposed to teach a computer to detect and prevent such attacks as intrusion into a network, stealing user credentials, moving between computers and, ultimately, stealing data. The majority of security systems in the present day operate based on a list of bad things. This cannot work when attackers switch techniques. The security teams are also overwhelmed with unnecessary alerts and are unable to react timely. The questionnaire includes the reasons why these outdated methods are no longer effective and what a more effective system may be like. The main components used are: a simulated real world that can be trained on, a cost-sensitive reward system, an intelligent adversary that learns and evolves, a human decision maker in control, and debugging tools to demonstrate why the AI made the choice it made. Three methods of reinforcement learning were compared. The one named PPO was the best since it learned more effortlessly and prevented more attacks compared to the others. What follows such as applying several AI agents collaboratively or integrating them with real security tools is also addressed in the paper. We discovered that it may actually work in reality to have a system learn by experience, particularly, when it is trained to respond as it does to a smart attacker. It learns how to trade off intelligently and does not always react in a similar manner.

Keywords: Advanced Persistent Threat (APT); Reinforcement Learning; Proximal Policy Optimization (PPO); Adversarial Machine Learning; Multi-Stage Cyber Defense; Human-in-the-Loop Security; Explainable Artificial Intelligence (XAI).

1. Introduction

These are silent and slow unlike the normal cyber attacks that strike fast and loud. Intruders do not destroy the front door. They come upon an open window somewhere. Perhaps, somebody pressed a button that is not supposed to be pressed [1]. Perhaps there is an old account that still exists and nobody has deleted it. When they are there they do not do anything blatant. They just sit and watch. They get to know how internally the network operates. Who talks to which servers. Hours of individual logging in and out. Which files are in fact used. It takes weeks or months before they are transferred to another machine and when this happens they snatch passwords in the process. No one knows that something is wrong before he or she has the keys to all that matters [2]. What is so difficult to prevent is how ordinary all their action appears to tools of security. They employ actual accounts in the process of doing business. They get into files that individuals are using. When they do it, they do not raise any alarm since they are not engaging in any activities that seem to be an attack [3]. Very few companies ever learn that they were hit until months later when a person finds their data somewhere being sold. Even at that time they may not have figured out how it came about [4].

The same tools that have been utilized most often by companies to defend themselves have failed in this type of attack [5]. Firewalls and antivirus softwares operate based on searching things that have already been known to be bad. They possess lists of attack signature and virus patterns. When they come across something that is similar they block. That is alright when it comes to attacks that happened a decade ago. But those tricks are not used by APTs. They switch it around so that they are able to get away with it each time. They could employ what regular system

administrators employ, and it may appear as maintenance as opposed to an attack [6]. Other systems attempt to identify odd behavior rather than attempting to identify familiar bad things. That is a smart idea in theory and creates enormous problems in reality. There is all sorts of strangeness in normal networks. A person enters in another floor. A server spikes for no reason. An employee logs in at 2am, as he/she was not able to sleep. None of that implies the presence of an attack [7]. These systems cause far too many false alarms. Thousands of alerts occur each day to security teams. You are not able to explore them all. After some time, people begin to disregard warnings completely since the majority of them are not something to worry about. The actual assault is lost in all the commotion [8]. And these systems examine alerts individually. They do not make the links between an event that occurred last Tuesday and the one that occurred three weeks later. Attackers count on that. They defer their activities over time in such a way that nothing will arouse a threshold [9].

Human aspect of security is frankly an shambles at present [10]. The security Operations Centers employ brains to sit in front of the screens observing threats. It is not something trivial but the real job is tiring. You look at dash boards with warnings. Most are false alarms. The ones that are not false alarms tend to be something stupid such as a person typing out their password incorrectly more than once. The hours pass and nothing real takes place. You get a sudden wake-up call then you are forced to bring yourself to complete attention [11]. Research indicates that the end of the shifts experience more threats than the beginning by analysts. Your brain simply cannot remain witty eight hours straight staring at notifications [12]. Even the good analysts who are really good, take time to research well. They are forced to plunge into logs, investigate systems, understand whether something really wrong is taking place. Attackers are not waiting when you do so. They keep moving. When one realizes that there is a real issue and they begin to block out things, the attackers are in another part of the network [13]. What is even worse is that companies understand that this is one of the issues but they just have no superior methods. There is no possibility of firing all of your analysts and hoping that things are all right. However, you can also not continue to add more people to the problem since the alerts are increasing at a faster rate than you have hired [14]. What is required is machines that can take the first level of response immediately. Human beings should leave the complex stuff that requires thinking. However, it is frightening to have the security decisions made independently by machines. In case they experience a mistake and block something important, it may bring major issues. Shutting down a server that a hospital needs can interfere with essential services and no one will want to be liable to such damages. [15].

This is why within a few years some researchers began to consider the reinforcement learning to provide cyber defense [16]. The fundamental concept is in fact a very simple one despite the complex mathematical analysis. You teach a computer program by trial and error, and it is like training a dog to do tricks using treats. It is rewarded when it makes an excellent defensive move. Bad moves get penalties. In time it will discover what works best and nobody need write the rules against everything that can happen [17]. Initial experiments demonstrated that this may be effective with basic things such as preventing one form of attack. But APTs are much more complex than that. They are long strands of little steps, which do not themselves imply much. It is not a big deal that one attempts to log in but fails. It does not matter whether one accesses a file. Hundred logins that failed on various accounts and then accessing files on a server which is not in use is a thing which is worth attention. [18]. You require a system which recognizes the entire sequence, and not just each piece alone. You must also have it to be concerned with cost.

Barring the whole may, technically, prevent attacks but it also prevents legitimate work. Isolate all the computers with a strange alert and no one gets any work done. That is not a viable solution to any company that attempts to remain in business [19]. The agent must acquire trade-offs. It is sometimes better to wait and see and stop not immediately. There are times when you have to make a move before it is too late. To know when to do what is the main trick [20].

More recent studies have developed real simulation space in which you can safely train these agents without destroying real company networks [21]. Whole attack campaigns with their twists and turns can be modeled. There is phishing, there is privilege escalation, everywhere is lateral. You can change the attacker to follow alternative routes so as to ensure that the defender does not simply memorize a single pattern [22]. Other teams also included smart attackers, which learn and evolve as well. That means that the defender must continue to improve just to match. When the attacker discovers that you always block one thing it will be some other thing. That compels the defender to study general tactics rather than mere rules [23]. Such adversarial training results in the final system that is far more difficult to overcome. It is as though preparing to battle by sparring someone who fights back rather than hitting a heavy bag [24]. The other researchers were engaged in making the decisions understandable to man. No one believes a computer that he or she cannot comprehend. You have to be able to know why it blocked that user. Why it flagged that file. Why it believed that server had to be isolated at this point in time [25]. Certain teams created dashboards that indicated what variables contributed to every choice. The remaining ones employed methods that estimate the importance of which information is important. This way an analyst can observe that the decision was informed by three factors, including weird login time, access to sensitive files, and unsuccessful attempts prior to that. That makes sense. That would have been looked at by a human too [26].

Another useful activity is the effort to put humans in the loop when it comes to the really big decisions [27]. All the little stuff is taken care of by the computer. Monitoring. Low-level blocks. Watching for patterns. However, it requests permission before it does something drastic such as isolating a whole server or terminating network accessibility to a whole department. A human receives a notification with the rationale and can either say yes or no to it in a time frame [28]. This preserves the advantage of speed offered by automation, and introduces the element of safety. When computer is not right on something big, one can notice before something goes wrong.

Unless the human is responsive and in time, the computer may default to a safer position or it may become aggressive to another individual [29].Such a hybrid solution as this is more realistic to real companies as compared to full automation. No one is willing to give away the keys. But everybody is fed up with being submerged in notifications as attacks continue to be successful. All that is considered in this survey. What has been experimented on by researchers of the last few years, we read. What really works and what is still theory. Which algorithms are the most suitable in various circumstances. How near we are to something that may in the real world assist real security teams in real work [30].

1.1. Study Objectives

- To get firsthand information on what was written about reinforcement learning to stop Advanced Persistent Threat attacks and what works and what does not in various research works.

- To draw parallels between the way various simulation environments are constructed to train the defense agents and determine what aspects really have to be decided to make the training realistic and good.
- To explore the way in which researchers model attackers that learn and change with time and examine the training of defenders to deal with constantly evolving strategies.
- To research the various methods individuals can incorporate cost awareness into their systems to train defenders how to prevent attacks without disrupting usual operations and squandering resources.
- To identify the research gaps that exist in the current literature and contemplate what remains to be done to get any of this to even benefit a real security team in the process of doing real work.

2. Literature Review

One of the oldest literature relating to this field examined power system and energy grids since attacks on these can have a real physical impact [1]. A 2026 study investigated the use of trust-aware federated deep reinforcement learning in energy markets. It was meant to allow various sections of the grid to learn without the exchange of sensitive information. This is important since attackers usually target energy system because they are aware that there can be no downtime. You cannot afford to close down and go to inquire. The researchers created a system, which would be able to identify suspicious behaviour and yet maintain customer information confidential. They applied it to the case of microgrids and discovered that it was more effective in the detection of attacks compared to the more outdated techniques [2]. Meanwhile, another community also published on the topic of multi-agent reinforcement learning to detect and track attacks [3]. They employed more than one learning agent as opposed to one. The various agents monitored various sections of the network and they exchanged their knowledge. This assisted in capturing attacks that are mobile since one agent would be able to see a small thing and then another would see something different and they both determine that it is an attack. The authors experimented on their system using simulated network attacks and discovered that it detected more threats as compared to single-agent systems [4]. The other team considered the development of energy systems in particular and employed physics-conscious adaptive control with reinforcement learning [5]. They created a system that knows how the physical components are expected to behave in order to detect when something is not normal even when the cyber cues appear to be normal. When one attempts to cause a freezer to gradually warm up over a few days to destroy medicine, the system could be detected as the temperature changes do not agree with physics as to how things should happen [6]. Another group considered the power systems in a different perspective using multi agent deep reinforcement learning in routing during an attack condition [7]. Power systems contain a lot of data flowing across them and attacks tend to attack the data flow. When the conventional routes were attacked, their system had alternative ways through which the data could be found. It could not allow communication to fail and instead it bypassed problems automatically. This is important since during an actual attack, you require portions of system to continue communicating with one another even at the time of being fired at [8].

Industrial systems were also the focus as the industries are the ones that run factories, power plants, and water treatment, etc. [9]. This is terrifying because they can cause bodily injuries when assaulted. The hierarchical

reinforcement learning was used in late 2025 in incident detection and response systems in industrial cyber-physical systems [10]. Hierarchical means that they decomposed the problem into strata. On lower layers, fast simple decisions were carried out. Bigger layers were those that related to bigger strategy. This enabled the system to react quickly to the blatant threats as well as to predict more in the long-term more complicated attacks. They tested it under simulated factory conditions and it was more successful at identifying the attack as compared to flat systems that tried to do everything at the same time [11]. The hierarchical approach also helped in the explainability component. You could know what layer had made a call and why in case something had not gone well, and it is important when you need the plant managers to trust the system [12]. The entire research direction of multi-step attack detection does not depend on reinforcement learning, however, it is an input to the RL-work [13]. There was a 2025 big review article, which vetted all the different techniques that individuals attempt to differentiate attacks that appear in multiple steps over time [14]. The authors have found out that most of the traditional tools remain ineffective in as far as linking events across the time. They receive one error of authorization one time and a mystic file access another time but never connect them. The methods of learning incorporated in the literature review were machines-based machine learning methods, graph-based methods and behavior modeling-based method. What was outstanding was the fact that there is no common style that suits everyone. Each has blind spots. The authors have claimed that the new systems should adopt multiple approaches instead of having all their eggs in one basket [15]. Another survey carried out approximately at the same time is the one of the electronic warfare and the cyberattacks on drones and UAVs [16]. It is different world but the patterns of attack are parallel to APTs. The drone attacks too are executed in stages. First they interfere with communications, then they deceive GPS and finally they get in control. Defenses and countermeasures were also assessed and some RL-based attacks also showed promise to adapt to new attack patterns in real time [17].

The agent-based systems were also observed in some unanticipated regions that pertain to the construction of autonomous defense mechanisms [18]. One of the papers of 2025 discussed FUKURO, an agent-based system used to control remote host resources on a real-time basis [19]. It was not so much about security as such but the technology is. The agents would view the behavior of the system and adjust to ensure that things go well. They were able to sense when something was amiss and act without the need of a human. The same fundamental methodology is used in security. You would like your agents to monitor what is going on and react immediately rather than be alerted after getting through a ticketing system [20]. The other study considered hybrid machine learning in detecting anomalies and predicting threats [21]. They also used a combination of several ML methods to intercept more attacks and reduce the false alarms. This is massive since it is the false alarms that kill the actual security tools. When your system cries wolf, it is ignored. The hybrid method was more effective in drawing a line between actual threats and typical weirdness when they tested network traffic data [22]. Other researchers stood aside and contemplated the reinforcement learning of proactive cybersecurity more broadly [23]. In one of the papers in 2025, various paradigms of how RL would contribute to dynamic risk management were defined [24]. The most important lesson was that it is impossible to simply react. You must predict the next action of attackers and prepare to counter the action. This implies the practice of how to learn things with time and knowing the objectives of attackers. Assuming they tend to escalate privileges once they have a foothold, you can be more alert of privilege

escalation indications rather than wait until privilege escalation occurs. In the paper, it was argued that RL is an appropriate choice in this regard given that it is intrinsically sensitive to sequences and reward in the future [25]. One more team examined machine learning in self-driving cars with the help of a scientometric review [26]. It is a fact that they reviewed publication tendencies rather than technology itself. Their discoveries are important though as they indicate the direction the field is taking. They discovered massive increases in the number of papers concerning security in the autonomous system, which explains why researchers regard this as a significant future issue requiring a solution [27].

In some of the recent work, knowledge graphs appeared as a means of getting familiar with APTs [28]. One of the papers presented an invention of what was named Trail, which is a knowledge graph based approach to attributing attacks to particular groups [29]. They constructed graphs related to the action of the attackers, tools, and the targets. This enabled them to observe trends in various attacks and determine who was likely to be the perpetrators. Defensively, this is critical to the extent that when you are aware of who is mounting an assault on you, then you can make an educated guess on the next line of attack, using the experience the aggressor had previously. There are various playbooks of different groups. Knowing the playbook will make you defend [30].

3. Problem Statement

Conventional security solutions such as firewall and antivirus software use signatures of known attacks implying that they entirely overlook any new threats that no one has ever encountered.

- Majority of the existing defense systems consider each alert individually rather than comparing the events of the past, and thus they do not notice multi-stage attacks where individual steps appear to be innocent.
- Thousands of false alarms come at security teams daily, and they do not have enough time to respond promptly, which gives attackers enough time to navigate networks before anyone realizes it.
- Current machine learning methods require enormous quantities of labeled history of attacks that are not available and once they learn they cannot update because the attackers evolve strategies.
- The study of reinforcement learning in cyber defense is largely done in simplified models that are not representative of the complexity of a real network, and the systems are not aware of the cost of operation or otherwise explain their actions to people who are studying them.

4. Conclusion

The reinforcement learning actually appears to have a chance of prevention of attacks by Advanced Persistent Threats, much higher than the old rule-based systems that everybody is still running. The articles that we have read demonstrate that good simulator trained agents can be taught to identify multi-stage attacks and respond before they get ugly. PPO continued to emerge first in other algorithms due to its ability to learn steadily and not to forsake old tricks in acquiring new ones. The key point is to create the appropriate simulation environment with realistic attacker behavior and cost sensitivity or the agent studies things that do not apply in the real world. The human piece is huge too. No one is going to allow a black box AI to issue large security calls without knowing the reason

so explainability tools must be included in any system that is actually used. Putting people in the loop when it comes to the big decision makes it relatively safe and leaves the automation to deal with all the small alerts that wear the analysts out. It is not done yet and no one of this will go to the real security teams. The majority of studies remain in the simulation and real networks are messier. Nonetheless, the field is converging to systems which learn, adapt and explain themselves as well as collaborate with people rather than attempting to eliminate them. It is as though that is the right direction.

Declarations

Source of Funding

This study did not receive any grant from funding agencies in the public, commercial, or not–for–profit sectors.

Competing Interests Statement

The authors have not declared any conflict of interest.

Consent for publication

The authors declare that they consented to the publication of this study.

Authors' contributions

All the authors equally contributed to this study.

Informed Consent

Not applicable for this study.

Institutional Review Board Statement

Not applicable for this study.

Ethical Approval

Not applicable for this study.

Declaration of Artificial Intelligence

The authors declare that no artificial intelligence (AI) tools or AI-assisted technologies were used in preparing this manuscript.

References

[1] Sepehrzad R., Hosseinalibeiki H., Taghinezhad N., Khosravi N., Al Durra A., & Sadabadi M.S. (2026). Enhanced cyber-resilience in flexible energy markets for microgrids: a trust-aware federated deep reinforcement learning framework. Applied Energy, 406: 127282. https://doi.org/10.1016/j.apenergy.2025.127282.

[2] Kalantri R.A., & Bansode R. (2026). Adaptive multiagent reinforcement learning framework for cyber attack detection and tracking. International Journal of Robust and Nonlinear Control. https://doi.org/10.1002/rnc.70415.

[3] Liu J., Wu B., Meng X., Wu J., & Ma Z. (2026). Physics-aware adaptive model predictive control guided by reinforcement learning for enhanced cyber-resilience of building energy systems. Energy Conversion and Management, 350: 120990. https://doi.org/10.1016/j.enconman.2025.120990.

[4] Xiahou K., Wang Z., Jiang D., Xu X., Zheng J., Liu Z., & Wu Q.H. (2026). Multi-agent deep reinforcement learning based robust routing strategy for cyber-physical power system considering cyber attacks. CSEE Journal of Power and Energy Systems. https://doi.org/10.17775/cseejpes.2025.05080.

[5] Babar A., Halabi T., & Zulkernine M. (2026). Autonomous and adaptive cyber incident detection and response in industrial cyber-physical systems using hierarchical reinforcement learning. ACM Transactions on Cyber-Physical Systems, 10(1): 1–27. https://doi.org/10.1145/3765622.

[6] Shaukat S.U., Khan S., & Parkinson S. (2025). A review on multi-step attack detection. IEEE Access. https://doi.org/10.1109/access.2025.3607497.

[7] Yu A., Kolotylo I., Hashim H.A., & Eltoukhy A.E.E. (2025). Electronic warfare cyberattacks, countermeasures and modern defensive strategies of UAV avionics: a survey. IEEE Access. https://doi.org/10.1109/access.2025.3561068.

[8] Abd Jalil K., Wei Yi K., & Naim M.H. (2025). FUKURO: an agent-driven framework for real-time remote host resource management. International Journal of Research and Innovation in Social Science, 9(10). https://doi.org/10.47772/ijriss.2025.910000556.

[9] Salman A.M., Al-Nuaimi B.T., Subhi A.A., Alkattan H., & Alfilh R.H.C. (2025). Enhancing cybersecurity with machine learning: a hybrid approach for anomaly detection and threat prediction. Mesopotamian Journal of CyberSecurity, 5(1): 202–215. https://doi.org/10.58496/mjcs/2025/014.

[10] Ren S., Chen S., & Zhang Q. (2025). Reinforcement learning paradigms for proactive cybersecurity and dynamic risk management. Frontiers in Artificial Intelligence Research, 2(3): 436–456. https://doi.org/10.71465/fair417.

[11] Hassan M., Kabir M.E., Islam M.K., Alam E., Rambe A.H., Jusoh M., & Sameer M. (2025). Mapping the machine learning landscape in autonomous vehicles: a scientometric review of research trends, applications, challenges and future directions. IEEE Access. https://doi.org/10.1109/access.2025.3620637.

[12] King I.J., Ramirez R., Bowman B., & Huang H.H. (2025). Trail: a knowledge graph-based approach for attributing advanced persistent threats. In Proceedings of the IEEE International Conference on Data Engineering, Pages 1207–1220. https://doi.org/10.1109/icde65448.2025.

[13] Meduri K., Gonaygunt H., & Nadella G.S. (2024). Evaluating the effectiveness of AI-driven frameworks in predicting and preventing cyber attacks. International Journal of Research Publication and Reviews, 5(3): 6591–6595. https://doi.org/10.55248/gengpi.5.0324.0875.

[14] Alnfiai M.M. (2025). AI-powered cyber resilience: a reinforcement learning approach for automated threat hunting in 5G networks. EURASIP Journal on Wireless Communications and Networking, 2025(1): 68. https://doi.org/10.1186/s13638-025-02497-2.

[15] Brandao P. (2025). Combating advanced persistent threats through artificial intelligence: an algorithmic approach. Open Research Europe, 5: 139. https://doi.org/10.12688/openreseurope.20268.2.

[16] Askhatuly A., Berdysheva D., Berdyshev A., Adamova A., & Yedilkhan D. (2025). Adversarial attacks and defense mechanisms in machine learning: a structured review of methods, domains and open challenges. IEEE Access. https://doi.org/10.1109/access.2025.3624409.

[17] Celdrán A.H., Sánchez Sánchez P.M., von der Assen J., Schenk T., Bovet G., Martínez Pérez G., & Stiller B. (2024). RL and fingerprinting to select moving target defense mechanisms for zero-day attacks in IoT. IEEE Transactions on Information Forensics and Security, 19: 5520–5529. https://doi.org/10.1109/tifs.2024.3402055.

[18] Osei A., Al Mtawa Y., & Halabi T. (2024). Mitigating adversarial reconnaissance in IoT anomaly detection systems: a moving target defense approach based on reinforcement learning. EAI Endorsed Transactions on Internet of Things, 10. https://doi.org/10.4108/eetiot.6574.

[19] Beltrán-López P., Gil Pérez M., & Nespoli P. (2025). Cyber deception: taxonomy, state of the art, frameworks, trends and open challenges. IEEE Communications Surveys & Tutorials. https://doi.org/10.1109/comst.2025.3594788.

[20] Zhang T., Au Yeung C., Aurelia E., Onishi Y., Chulpongsatorn N., Li J., & Tang A. (2025). Prompting an embodied AI agent: how embodiment and multimodal signaling affects prompting behaviour. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Pages 1–25. https://doi.org/10.1145/3706598.3713110.

[21] Siddiqui E.F., Haleem M., Ahmad S.F., Salhi A., Zamani A.T., & Varish N. (2025). A multi-layered AI-driven cybersecurity architecture: integrating entropy analytics, fuzzy reasoning, game theory and multi-agent reinforcement learning for adaptive threat defense. IEEE Access. https://doi.org/10.1109/access.2025.3610526.

[22] Latif R.M.A., Ullah F., Raza U., Mostarda L., Khan M.A., & Cacciagrano D. (2025). Real-time threat detection using stream analytics and deep learning on network logs. In Proceedings of the IEEE International Conference on Big Data Science and Engineering, Pages 1–5. https://doi.org/10.1109/icbdse65491.2025.11220098.

[23] Malik J., Muthalagu R., & Pawar P.M. (2024). A systematic review of adversarial machine learning attacks, defensive controls and technologies. IEEE Access, 12: 99382–99421. https://doi.org/10.48550/arxiv.2509.20411.

[24] Sammartino V. (2025). A framework for proactive cyber-resilience: non-intrusive modeling for autonomous defense. In Proceedings of the International Symposium on Distributed Simulation and Real Time Applications, Pages 1–4. https://doi.org/10.1109/ds-rt68115.2025.11185079.

[25] Chao P., Robey A., Dobriban E., Hassani H., Pappas G.J., & Wong E. (2025). Jailbreaking black box large language models in twenty queries. In Proceedings of the IEEE Conference on Secure and Trustworthy Machine Learning, Pages 23–42. https://doi.org/10.1109/satml64287.2025.00010.

[26] Cao Y., Mahanti A., & Naha R. (2025). A real-time defense framework using PPPO in deep reinforcement learning for CyberBattleSim. In Proceedings of the IEEE International Conference on Trust, Security and Privacy in Computing and Communications, Pages 3024–3031. https://doi.org/10.1109/Trustcom66490.2025.00360.

[27] Lallie H.S., Thompson A., Titis E., & Stephens P. (2025). Analysing cyber attacks and cyber security vulnerabilities in the university sector. Computers, 14(2): 49. https://doi.org/10.3390/computers14020049.

[28] Anand N., Saifulla M.A., Ponnuru R.B., Alavalapati G.R., Patan R., & Gandomi A.H. (2024). Securing software defined networks: a comprehensive analysis of approaches, applications and future strategies against DoS attacks. IEEE Access. https://doi.org/10.1109/access.2024.3520478.

[29] He J., Zhao D., Liu T., Zou Q., & Xie J. (2025). Research on adaptive reward optimization method for robot navigation in complex dynamic environment. Computers, Materials & Continua, 84(2). https://doi.org/10.32604/cmc.2025.065205.

[30] Lee S., Lee K., Cho S., & Choi C. (2025). APTStop: a real-time framework for APT defense via strategic threat observation and prediction. IEEE Access. https://doi.org/10.1109/access.2025.3624035.

[31] Ronanki S.K., & Killi B.R. (2026). Machine learning based cyber attack detection and classification in O-RAN. Journal of Network and Systems Management, 34(2): 41. https://doi.org/10.1007/s10922-026-10035-8.

[32] Guntupalli R. (2024). Enhancing cloud security with AI: a deep learning approach to identify and prevent cyberattacks in multi-tenant environments. SSRN. https://doi.org/10.4018/979-8-3693-2639-8.ch011.

[33] Mahamkali N., & Mudigonda K.S.P. (2026). An ensemble framework for effective detection and classification of cyber attacks using machine learning. In AI Solutions for Detecting Cyber-Attacks in Information Systems, Pages 65–96. https://doi.org/10.4018/979-8-3373-1807-3.ch004.

[34] Ajala O.A., Okoye C.C., Ofodile O.C., Arinze C.A., & Daraojimba O.D. (2024). Review of AI and machine learning applications to predict and thwart cyber-attacks in real-time. Magna Scientia Advanced Research and Reviews, 10(1): 312–320. https://doi.org/10.30574/msarr.2024.10.1.0037.

[35] Ren S., Jin J., Niu G., & Liu Y. (2025). ARCS: adaptive reinforcement learning framework for automated cybersecurity incident response strategy optimization. Applied Sciences, 15(2): 951. https://doi.org/10.3390/app15020951.

[36] Khalaf M.A., & Steiti A. (2024). Artificial intelligence predictions in cyber security: analysis and early detection of cyber attacks. Babylonian Journal of Machine Learning, 2024: 63–68. https://doi.org/10.58496/bjml/2024/006.

[37] Manikandan A., & Deepa Rajan S. (2025). Cyber attack detection using deep multi-agent reinforcement learning with Beth dataset. SN Computer Science, 6(5): 433. https://doi.org/10.1007/s42979-025-03981-8.

[38] Al-Nawashi M.M., Al-hazaimeh O.M., Tahat N.M., Gharaibeh N., Abu-Ain W.A., & Abu-Ain T. (2025). Deep reinforcement learning-based framework for enhancing cybersecurity. International Journal of Interactive Mobile Technologies, 19(3). https://doi.org/10.3991/ijim.v19i03.50727.

[39] Hossain M.A. (2025). Deep Q-learning intrusion detection system (DQ-IDS): a novel reinforcement learning approach for adaptive and self-learning cybersecurity. ICT Express. https://doi.org/10.1016/j.icte.2025.05.007.

[40] Emirmahmutoğlu E., & Atay Y. (2025). A feature selection-driven machine learning framework for anomaly-based intrusion detection systems. Peer-to-Peer Networking and Applications, 18(3): 161. https://doi.org/10.1007/s12083-025-01947-4.

[41] Yang W., Acuto A., Zhou Y., & Wojtczak D. (2026). A survey for deep reinforcement learning based network intrusion detection. Applied AI Letters, 7(2): e70026. https://doi.org/10.1002/ail2.70026.

[42] Saeed M.M. (2025). An AI-driven cybersecurity framework for IoT: integrating LSTM-based anomaly detection, reinforcement learning and post-quantum encryption. IEEE Access. https://doi.org/10.1109/access.2025.3576506.

[43] Vladov S., Jotsov V., Sachenko A., Prokudin O., Ostapiuk A., & Vysotska V. (2025). Neural network method of analysing sensor data to prevent illegal cyberattacks. Sensors, 25(17): 5235. https://doi.org/10.3390/s25175235.

[44] Kharra S., & Chaudhary U. (2025). A novel approach of ontology-based cyber attack detection and prevention system using AI. In Proceedings of the International Conference on Computing Technologies & Data Communication, Pages 1–7. https://doi.org/10.1109/icctdc64446.2025.11158769.

[45] Alohali M.A., Dafaalla H., Baihan M., Alahmari S., Ben Miled A., Alrusaini O., Alqazzaz A., & Alkhudhayr H. (2025). Leveraging self attention driven gated recurrent unit with crocodile optimization algorithm for cyberattack detection using federated learning framework. Scientific Reports, 15(1): 23805. https://doi.org/10.1038/s41598-025-99452-4.

[46] Chinnasamy P., Yarramsetti S., Ayyasamy R.K., Rajesh E., Vijayasaro V., Pandey D., Pandey B.K., & Lelish M.E. (2025). AI-driven intrusion detection and prevention systems to safeguard 6G networks from cyber threats. Scientific Reports, 15(1): 37901. https://doi.org/10.1038/s41598-025-21648-5.

[47] Gujar S.S. (2024). Optimizing threat mitigation in critical infrastructure through AI-driven cybersecurity solutions. In Proceedings of the Global Conference on Communications and Information Technologies, Pages 1–7. https://doi.org/10.1109/gccit63234.2024.10862689.

[48] Ankalaki S., Atmakuri A.R., Pallavi M., Hukkeri G.S., Jan T., & Naik G.R. (2025). Cyber attack prediction: from traditional machine learning to generative artificial intelligence. IEEE Access, 13: 44662–44706. https://doi.org/10.1109/access.2025.3547433.

[49] Singh T. (2025). Artificial intelligence-driven cyberattacks. In Cybersecurity, Psychology and People Hacking, Pages 167–188. https://doi.org/10.1007/978-3-031-85994-6_17.

[50] Alomiri A., Mishra S., & AlShehri M. (2024). Machine learning-based security mechanism to detect and prevent cyber-attack in IoT networks. International Journal of Computing and Digital Systems, 16(1): 645–659. https://doi.org/10.12785/ijcds/160148.